Los Alamos
NATIONAL LABORATORY
research draft

Applied Theoretical & Computational Physics
Transport Methods Group
http://www-xdiv.lanl.gov/XTM/

To/MS: NS&E Draft Technical Note
From/MS: H. Lichtenstein (X-TM)/D409
Phone/FAX: 7–3926/5–5538
Email: hal@lanl.gov
Symbol: LA-UR-98-5715
Date: December 9, 1998

Exponential Convergence Rates for Reduced-Source Monte Carlo Transport in $[x, \mu]$ Geometry

Los Alamos National Laboratory

Abstract

This technical note describes a utilization, in an adaptive reduced-source approach, of a Monte Carlo transport solution for the one-speed finite slab problem in $[x, \mu]$ geometry.[1] Although a solution for the underlying problem has been available to arbitrary precision for some time, the purpose here is to demonstrate how the convergence afforded by traditional (non-adaptive) Monte Carlo can be improved significantly, without compromising its precision. It is demonstrated that the reduced-source Monte Carlo technique obtains multiple orders-of-magnitude improvement over traditional Monte Carlo convergence, for the two-dimensional transport problem treated. The goal is that ongoing research will obtain exponential convergence for practical applications that are not tractable with methodology currently available.

1. Introduction

The purpose of the effort documented herein is to demonstrate the possibility of far exceeding the convergence efficiency afforded by traditional (*i.e.*, non-adaptive) Monte Carlo. The reported findings demonstrate exponential convergence for a transport solution of a two-dimensional problem, specifically, the one-speed homogeneous slab with isotropic scattering and boundary sources. It is common knowledge that a solution to such a problem has been available to arbitrary precision for some time. The interest here, however, is *not* to offer an alternative solution technique for this problem, per se, but to exploit the convenience of a known solution, together with the availability of a recently published Monte Carlo approach for it,[1] in order to demonstrate the vast potential of adaptive Monte Carlo through the reduced-source method.

The goal is that ongoing research will enable the extension of adaptive Monte Carlo techniques to the point where practical applications, for which sufficiently dextrous methodology is currently unavailable, can be tackled efficiently. The present findings are part of ongoing, inter-division research at Los Alamos National Laboratory,[2] which has the ambition of advancing the applicability of adaptive Monte Carlo techniques. My current findings derive from an adaptation of an aforementioned work,[1] using it as the kernel within a reduced-source approach, whence derives the exponential convergence to the exact solution. The exponential convergence in the present results is shown to exceed by as much as ten orders of magnitude the convergence effected by the inverse root of computing time of non-adaptive Monte Carlo.

2. Preliminaries

In order to focus attention on that which is new in the present technical note, the starting point for this research is presented in this section with a minimum of discussion. That is, Equations 1–10, below, are presented for convenience, but are neither derived nor discussed herein, except for some basic definitions of terms.

Case's method,[3] following the notation in Bell and Glasstone,[4] gives the analytic form of the transport solution to the one-speed finite slab problem (with isotropic scattering) in $[x, \mu]$ geometry, *viz.*

$$\Phi(x,\mu) = a_+\psi_0^+(\mu)e^{-x/\nu_0} + a_-\psi_0^-(\mu)e^{x/\nu_0} + \int_{-1}^{1} A(\nu)\psi_\nu(\mu)e^{-x/\nu}d\nu \tag{1}$$

where x is the position (measured along the slab normal); μ is the direction cosine with respect to the x-axis; and $\Phi(x,\mu)$ is the angular flux.

The $\pm\nu_0$ are the discrete eigenvalues of the associated eigenfunctions

$$\psi_0^{\pm}(\mu) = \frac{c}{2}\frac{\nu_0}{\nu_0 \mp \mu} \tag{2}$$

with ν_0 being the positive root of

$$1 = c\nu_0 tanh^{-1}\frac{1}{\nu_0} \equiv \frac{c\nu_0}{2} ln\frac{\nu_0+1}{\nu_0-1} \tag{3}$$

The ν are in the continuum of eigenvalues, corresponding to the eigenfunction continuum

$$\psi_\nu(\mu) = \frac{c}{2}P\frac{\nu}{\nu-\mu} + \lambda(\nu)\delta(\mu-\nu) \tag{4}$$

with

$$\lambda(\nu) = 1 - \frac{c\nu}{2} ln\Big(\frac{1+\nu}{1-\nu}\Big) \tag{5}$$

where c is (for non-multiplying media) the collision survival probability; P indicates that the Cauchy principal value is to be used in any integration of a singular term; and $\delta(x)$ is the Dirac delta function. Orthogonality conditions on $\psi_0^{\pm}(\mu)$ and $\psi_\nu(\mu)$ lead to

$$a_{\pm} = \frac{1}{N_0^{\pm}} \int_{-1}^{1} \Phi(0,\mu)\mu\psi_0^{\pm}(\mu)d\mu \tag{6}$$

$$A(\nu) = \frac{1}{N_\nu} \int_{-1}^{1} \Phi(0,\mu)\mu\psi_\nu(\mu)d\mu \tag{7}$$

$$N_0^{\pm} = \int_{-1}^{1} \mu\psi_0^{\pm}(\mu)\psi_0^{\pm}(\mu)d\mu = \pm\frac{c}{2}\nu_0^3\Big[\frac{c}{\nu_0^2-1} - \frac{1}{\nu_0^2}\Big] \tag{8}$$

$$\int_{-1}^{1} \mu\psi_{\nu'}(\mu)\psi_\nu(\mu)d\mu = N_\nu\delta(\nu-\nu') \tag{9}$$

$$N_\nu = \nu\Big[\lambda^2(\nu) + \frac{\pi^2c^2}{4}\nu^2\Big] \tag{10}$$

Based on the foregoing, Booth [1] derived and prescribed Monte Carlo estimation for quantities that are used to obtain the coefficients $A(\nu)$ in Equation 7. These, together with the other computed quantities specified in Equations 2–10, are used to obtain an estimate of the flux specified in Equation 1. Such an estimation of the angular flux has been adapted as the kernel computation for the iterative approach described in the following sections.

3. The Reduced Source at the Boundary

If the slab thickness is T, then the left and right boundaries may be specified as $x = 0$ and $x = T$, or, equivalently, as $x = -T$ and $x = 0$, respectively. The first specification is convenient for left-boundary computations; the second for right boundary computations.

Given the computed angular flux, $\Phi(x,\mu)$, evaluated at the slab boundaries (initially choosing $x = 0$ to be the left boundary and $x = T$ the right), the corresponding (inward directed) boundary-source density is given by

$$S(\mu) = \mid \mu \mid [\Phi(0,\mu)H(\mu) + \Phi(T,\mu)H(-\mu)] \tag{11}$$

where $H(z) = \int_{-\infty}^{z} \delta(y)dy$ is the Heaviside function, and $\Phi(x,\mu)$ is given by Equation 1.

If the initial computed solution of $\Phi(x,\mu)$, say $\Phi^0(x,\mu)$, is equal to the exact solution for the given problem, say $\Phi^E(x,\mu)$, then $S(\mu)$ will be exactly equal to the given source at the boundary, say $S^E(\mu)$.

Any non-vanishing *algebraic* difference between the given source, $S^E(\mu)$, and the computed source, from Equation 11, comprises the reduced source at the boundary, or boundary residual,

$$S^1(\mu) = S^E(\mu) - S(\mu) \tag{12}$$

$S^1(\mu)$, in turn, specifies the source for the problem, whose solution, say $\Phi^1(x,\mu)$, is the difference between the exact solution to the original problem and the solution given by $\Phi^0(x,\mu)$, *viz.*

$$\Phi^E(x,\mu) = \Phi^0(x,\mu) + \Phi^1(x,\mu) \tag{13}$$

It should be noted that $S^1(\mu)$ can, and in practice does, have both positive and negative components. This is treated by assigning positive and negative weights, respectively, in the Monte Carlo estimation.

Now, if $\Phi^1(x,\mu)$ is not an *exact* solution (to the *residual* problem for which $S^1(\mu)$ is the given source), as might be evidenced by some *sufficiently* large boundary residual, say $S^2(\mu)$, the process can be continued for another iteration. The choice of a termination criterion, based on a specified magnitude of boundary residual, or some more appropriate metric (see Section 7, below), is arbitrary.

In this manner, Equations 11–13 prescribe an iterative method for converging to the exact solution of the original problem — the reduced-source method.

4. Direct Contributions to Transport Coefficients from Boundary Residuals

The kernel for this reduced-source iteration accounts for the estimation of the transport coefficients, except for the *direct* contribution from the boundary residual of Equation 12. That direct contribution can be computed by numerical integration, analogous to the treatment in Reference [1].

Subscripting with L to indicate left-boundary quantities, the left-boundary flux due to the left-boundary residual source density is, from Equations 11 and 12

$$\phi_L(\mu)H(\mu) = \frac{S_L^1(\mu)}{|\mu|} = \frac{S_L^E(\mu)}{|\mu|} - \Phi(0,\mu)H(\mu) \tag{14}$$

where $S_L^E(\mu)$ is the given source (on the left boundary) and $\Phi(0,\mu)$ is the computed flux (on the left boundary).

Substituting Equation 14 into Equation 7 we get

$$A_L(\nu) = \frac{1}{N_\nu}\int_{-1}^{1} \phi_L(\mu)H(\mu)\mu\psi_\nu(\mu)d\mu \tag{15}$$

Using Equations 4 and 15 we get

$$\begin{aligned} A_L(\nu) &= \frac{1}{N_\nu}\int_{-1}^{1} \phi_L(\mu)H(\mu)\mu\Big[\frac{c}{2}P\frac{\nu}{\nu-\mu} + \lambda(\nu)\delta(\mu-\nu)\Big]d\mu \\ &= \frac{1}{N_\nu}\int_{0}^{1} \phi_L(\mu)\mu\frac{c}{2}P\frac{\nu}{\nu-\mu}d\mu + \frac{1}{N_\nu}\phi_L(\nu)H(\nu)\nu\lambda(\nu) \\ &= I_L(\nu) + B_L(\nu) \end{aligned} \tag{16}$$

where

$$I_L(\nu) = \frac{1}{N_\nu}\int_{0}^{1} \phi_L(\mu)\mu\frac{c}{2}P\frac{\nu}{\nu-\mu}d\mu \tag{17}$$

and

$$B_L(\nu) = \frac{1}{N_\nu}\phi_L(\nu)H(\nu)\nu\lambda(\nu) \tag{18}$$

Using Equations 10 and 18 we get

$$\begin{aligned} B_L(\nu) &= \frac{\phi_L(\nu)H(\nu)\nu\lambda(\nu)}{\nu\Big[\lambda^2(\nu) + \frac{\pi^2c^2}{4}\nu^2\Big]} \\ &= \frac{\phi_L(\nu)H(\nu)}{\Big[\lambda(\nu) + \frac{\pi^2c^2}{4\lambda(\nu)}\nu^2\Big]} \end{aligned} \tag{19}$$

From Equation 5, we see that $\lambda(\nu)$ is unbounded for $\nu = \pm 1$; but this presents no difficulty, since $B_L(\pm 1)$ vanishes, as can be seen from Equation 19.

The $I_L(\nu)$ of Equation 17 can be calculated, with appropriate care in the vicinity of $\nu = \mu$, using the approach in Reference [1], Section VII, and also noting from Reference [1], Section IV, that we need $A_L(\nu)$ only for $\nu > 0$.

Divide the interval [-1,1] into $2K$ equal intervals, and, for $0 \leq \nu \leq 1$, let ν_i and μ_i be the midpoints of the intervals

$$\mu_i = \nu_i = \frac{(i - \frac{1}{2}) - K}{K} \qquad i = K+1, K+2, \cdots, 2K \tag{20}$$

Let

$$l_i = \frac{i - K}{K} \qquad i = K+1, K+2, \cdots, 2K \tag{21}$$

be the endpoints of the i^{th} interval, *viz.*

$$l_{i-1} \leq \nu < l_i \qquad i = K+1, K+2, \cdots, 2K \tag{22}$$

To evaluate $I_L(\nu)$ for the eigenvalues ν_j, assume a linear form for

$$\alpha(\mu) = \frac{1}{2N_\nu}\phi_L(\mu)c\mu\nu \tag{23}$$

in each interval. That is,

$$\alpha_i(\mu) = a_i(\nu_j - \mu) + c_i \qquad i = K+1, K+2, \cdots, 2K \tag{24}$$

Note that a_i in Equation 24 is the slope of our linearized $\alpha_i(\mu)$ in each interval. For $i = K+1$ and $i = 2K$ we use a one-sided estimate of the a_i, because no information about $\phi_L(\mu)H(\mu)$ exists for $i = K$ or $i = 2K+1$. Elsewhere, a two-sided estimate of the slope is used.

Hence

$$a_{K+1} = \frac{\alpha(\mu_{K+2}) - \alpha(\mu_{K+1})}{\mu_{K+1} - \mu_{K+2}} = K[\alpha(\mu_{K+1}) - \alpha(\mu_{K+2})] \tag{25}$$

$$c_{K+1} = \alpha(\mu_{K+1}) - a_{K+1}(\nu_j - \mu_{K+1}) \tag{26}$$

$$a_i = \frac{\alpha(\mu_{i+1}) - \alpha(\mu_{i-1})}{\mu_{i-1} - \mu_{i+1}} = \frac{K}{2}[\alpha(\mu_{i-1}) - \alpha(\mu_{i+1})] \qquad i = K+2, K+3, \cdots, 2K-1 \tag{27}$$

$$c_i = \alpha(\mu_i) - a_i(\nu_j - \mu_i) \qquad i = K+2, K+3, \cdots, 2K-1 \tag{28}$$

$$a_{2K} = \frac{\alpha(\mu_{2K}) - \alpha(\mu_{2K-1})}{\mu_{2K-1} - \mu_{2K}} = K[\alpha(\mu_{2K-1}) - \alpha(\mu_{2K})] \tag{29}$$

$$c_{2K} = \alpha(\mu_{2K}) - a_{2K}(\nu_j - \mu_{2K}) \tag{30}$$

Thus, in the i^{th} interval, we get from Equation 17 and the definitions 20–30

$$\begin{aligned} I_L^i(\nu_j) &= \int_{l_{i-1}}^{l_i} \frac{a_i(\nu_j - \mu) + c_i}{\nu_j - \mu} d\mu \\ &= \frac{a_i}{K} + c_i ln\left|\frac{\nu_j - l_{i-1}}{\nu_j - l_i}\right| \end{aligned} \tag{31}$$

And note that for the case of $i = j$ we have

$$ln\left|\frac{\nu_j - l_{i-1}}{\nu_j - l_i}\right| = ln\left|\frac{\nu_j - l_{j-1}}{\nu_j - l_j}\right| = ln\left|\frac{\frac{1}{2K}}{\frac{-1}{2K}}\right| = ln(1) = 0 \tag{32}$$

so that

$$I_L^j(\nu_j) = \frac{a_j}{K} \tag{33}$$

Hence, the value of $I_L(\nu)$ in Equation 17 can be approximated for $\nu = \nu_j$ such that

$$I_L(\nu) \simeq \sum_{i=K+1}^{2K} I_L^i(\nu_j) \tag{34}$$

All the other required direct contributions are obtained in a similar fashion, yielding:

$$a_+ \simeq \sum_{i=K+1}^{2K} a_+^i \tag{35}$$

where

$$\begin{aligned} a_+^i &= \int_{l_{i-1}}^{l_i} \frac{a_i(\nu_0 - \mu) + c_i}{\nu_0 - \mu} d\mu \\ &= \frac{a_i}{K} + c_i ln\left(\frac{\nu_0 - l_{i-1}}{\nu_0 - l_i}\right); \end{aligned} \tag{36}$$

$$A_R(\nu) \simeq \sum_{i=1}^{K} I_R^i(\nu_j) + \frac{\phi_R(\nu) H(-\nu)}{\left[\lambda(\nu) + \frac{\pi^2 c^2}{4\lambda(\nu)}\nu^2\right]} \tag{37}$$

where

$$\begin{aligned} I_R^i(\nu_j) &= \int_{l_{i-1}}^{l_i} \frac{a_i(\nu_j - \mu) + c_i}{\nu_j - \mu} d\mu \\ &= \frac{a_i}{K} + c_i ln\left|\frac{\nu_j - l_{i-1}}{\nu_j - l_i}\right|; \end{aligned} \tag{38}$$

and finally

$$a_- \simeq \sum_{i=1}^{K} a_-^i \tag{39}$$

where

$$\begin{aligned} a_-^i &= \int_{l_{i-1}}^{l_i} \frac{a_i(\nu_0 + \mu) + c_i}{\nu_0 + \mu} d\mu \\ &= \frac{a_i}{K} + c_i ln\left(\frac{\nu_0 + l_i}{\nu_0 + l_{i-1}}\right) \end{aligned} \tag{40}$$

The omitted details may be found in a Los Alamos National Laboratory report.[5]

5. Implementation of Reduced Source Algorithms into Test-bed Code

The implementation of the algorithms implied by the foregoing sections is straightforward to describe. What is required, essentially, is:

1. an iteration loop surrounding the code kernel comprising Booth's [1] Monte Carlo estimation (designed by him for one pass of a specified number of histories);
2. the coding that computes the reduced source at the end of each iteration (see Section 3, above);
3. coding to "group-by-sign" and sample the reduced source (which, in general, will have both positive and negative components on both sides of the slab);
4. and, coding that computes, by numerical integration, the *direct* contributions to the transport coefficients from the reduced source (see Section 4, above).

Details of the implementation, including a test enabled by the odd-function property (with respect to the source) of the solution, may be found in a Los Alamos National Laboratory report.[5]

6. Demonstration of Reduced-Source Behavior

Example behavior of the boundary-source residuals is shown in Figures 1 and 2, below. Figure 1 displays the behavior of the positive components of the boundary-source residuals; Figure 2 displays the corresponding behavior of the negative components. As pointed out in Section 3, above, the boundary-source residuals generally have both positive and negative components.

Each of these figures compares the convergence of: one run that used 1000 bins for the numerical integrations over the eigenvalues ν (see Equation 1), and the same number for sampling the direction cosines μ from the reduced sources (see Equation 12); and, one run that used twice as many bins. In addition, a plot of a traditional inverse-root convergence rate is displayed for illustration. And, to highlight the comparison between the '1000-bin' result and the '2000-bin' result, the 1000-bin result is additionally shown scaled down to match the 2000-bin result at the third iteration. This serves to demonstrate that the convergence rate toward a zero boundary-source residual is constrained by the precision of the numerical computation. It is important to point out that central processing unit (CPU) time per iteration is roughly constant in a given run, but the magnitude of the constant, of course, depends on the precision of the numerical integration and source-sampling tables (as specified by the refinement of the binning).

7. Metric for Gauging Convergence of the Angular Flux Solutions

Although the boundary-source residuals appear to converge to zero (see Figures 1 and 2), it was seen upon reflection, that they need not converge to zero (let alone at a rapid rate), in order for the *flux solution* to converge toward the exact solution. This may be clarified by considering the consequences of the following behavior for the boundary-source residuals:

> Suppose the boundary-source residuals were characterized by an increase, per iteration, in the fluctuation frequency (about zero) of a non-vanishing amplitude, as a function of angle cosine.

It seems plausible that the net flux effected by such a non-vanishing boundary-source residual would tend to zero, by the cancellation effect of neighboring positive and negative source components. Hence, another computed quantity was needed as an appropriate metric with which to analyze convergence rates for the flux solutions of the reduced-source method.

In order to have the computed angular fluxes converge to the exact solutions, some aspect of a reduced-source iteration would have to become vanishingly small. Clearly, if the computed fluxes were the exact solutions, any subsequent iteration would be based on a boundary-source residual that was *effectively* zero (refer to Equation 12 and to the *gedankenexperiment*, above). Such a residual source would, necessarily, effect a vanishingly small residual scalar flux at the boundary. Hence, the residual scalar flux is a reasonable candidate for gauging the convergence of the angular flux solutions by the reduced-source method.

8. Anticipated Convergence Rates

The reduced-source method is expected to produce a rate of convergence (to the exact solution) substantially greater than that offered by the traditional Monte Carlo transport method.[7] Traditional Monte Carlo exhibits the $\sim \frac{1}{\sqrt{N}}$ (*i.e.*, inverse root) rate, where N is the number of histories sampled (or, equivalently, a number of CPU-time units). The reduced-source method is believed to offer exponential rates of convergence: $\sim (constant)^{-N}$, where, again, N has the dimension of some measure of computing time, as, for example, an iteration of the reduced-source method (given that each iteration takes roughly an equal amount of computing time for a specific problem). Such a rate of convergence would manifest itself graphically as a straight line, with negative slope, for a semilogarithmic plot of convergence metric as a function of iteration.

9. Convergence of the Residual Scalar Flux

Figures 3 and 4 demonstrate how the magnitude of the residual scalar flux converges as a function of iteration number. The residual scalar flux is obtained by numerical integration of the residual angular flux; and the *magnitude* is plotted, because the source residuals comprise, in general, both positive and negative components (thereby yielding both positive and negative flux residuals).

Figure 3 displays the behavior of the scalar flux residuals for the 5 thicknesses specified in its corresponding figure caption (see the section FIGURE CAPTIONS, below) on the left boundary of the slabs (*i.e.*, at the given source plane of incidence, $x = 0$). Figure 4 displays the corresponding behavior on the right boundary (at $x = T$). Both figures (for all 5 thicknesses) display two basic slopes: a steep slope for the first few iterations, followed by a shallower slope for later iterations.

In trying to understand this prominent feature of both figures, it was speculated [6] that the observed separation (between early and late iterations) of *relatively* 'fast' and 'slow' convergence rate was somehow effected by the two distinct natures of the components of the solutions (see Equation 1): the presumed dominance over the residual source, in the early iterations, of the asymptotic solutions (discrete eigenfunction components), and the later dominance of the transient solution (given by the integral over the continuum of eigenfunctions). Though seemingly a plausible conjecture, this was demonstrated by subsequent calculations (discussed below) to be off the mark.

Figure 5 demonstrates how the residual scalar flux converges both at the slab boundaries and at the midpoint of the slab. Here the slab thickness was taken to be 40 mfp and the given source at $x = 0$ was specified as before: $S(0, \mu) = \mu e^{-\mu^2}$. The observed convergence behavior of the plotted residual-flux metric as a function of iteration number is similar to that observed in Figures 3 and 4.

Since Figure 5 merely supports the observations from Figures 3 and 4, it is opportune, at this point, to remark upon the displayed convergence rate for the residuals, as compared to the familiar inverse-root rate, which is included for illustration. The inverse-root convergence illustration is seen to diminish by a factor of 5 in 25 iterations (by definition). This compares with the following decreases for the reduced-source metrics: more than 11 orders of magnitude at $x = 0$; and more than 10 orders of magnitude both at $x = 20$ and at $x = 40$.

10. Efficiency of Calculations

Exponential convergence, as described in the preceding section, is clearly desirable. But it is useful to examine how these convergence rates translate into traditional efficiency metrics.

I proceeded to calculate the problem defined as follows: a 1-mfp-thick slab in $[x, \mu]$ geometry, having a given source at $x = 0$, *viz.*

$$S(0, \mu) = \mu \qquad \mu > 0 \tag{41}$$

for which the analytic (inward-directed) boundary angular flux is:

$$\Phi(0, \mu) = 1 \qquad \mu > 0 \tag{42}$$

I timed the calculation and found that 10^6 histories took about 5 times as long as 100 histories. This may be surprising to those accustomed to traditional Monte Carlo transport, but the random walk in the test-bed code is a very small part of the CPU time per iteration. Most of the computing involves numerical integrations over the eigenvalues ν for the continuum of Case eigenfunctions (see, for example, Equation 1, above).

I then compared 1 iteration of 10^6 histories with 5 iterations of 100 histories/iteration. I also compared 3 iterations of 10^6 histories/iteration with 15 iterations of 100 histories/iteration. In so doing, I approximately equalized the total CPU time for the corresponding cases to be compared. Note, however, that the 10^6 histories exceeded the $\sim$*CPU-equivalent* 500 histories by a factor of 2×10^3.

TABLE I
Comparison of 1 Iteration (10^6 histories)
with 5 Iterations (10^2 histories/iteration)

(computed flux shown is for μ-bin having maximum deviation from 1)						
#iter	hist/iter	CPU (sec)	flux	deviation	(deviation)^2	rel FOM
1	10^6	184	1.0264	2.64×10^{-2}	6.97×10^{-4}	1
5	10^2	179	1.000338	3.38×10^{-4}	1.14×10^{-7}	6285

TABLE II
Comparison of 3 Iterations (10^6 histories/iteration)
with 15 Iterations (10^2 histories/iteration)

(computed flux shown is for μ-bin having maximum deviation from 1)						
#iter	hist/iter	CPU (sec)	flux	deviation	(deviation)^2	rel FOM
3	10^6	554	1.00153	1.53×10^{-3}	2.34×10^{-6}	1
15	10^2	535	1.0000135	1.35×10^{-5}	1.82×10^{-10}	13314

As part of the problem definition, I specified 1000 equal μ-bins for the range $-1 \leq \mu \leq 1$. And in Tables I and II above, the computed inward-directed angular flux quoted is for the bin value that deviated most from the *known* analytic value 1. The figure of merit (FOM) is computed as $((deviation)^2 \times CPU)^{-1}$ and is normalized to the smaller value of the two corresponding cases. A comparison of the relative FOM values is another indication that the relative improvement in efficiency provided by the reduced-source method involves multiple orders of magnitude.

11. Analysis of Convergence Rates for Flux Components

In order to investigate the conjecture concerning the two distinct rates of convergence evidenced in Figures 3–5 (see Section 9, above), I analyzed the three separable contributions to the angular flux: the two components comprising the asymptotic portion (as denoted by their associated coefficients $a_{\pm}$), and the transient component (denoted by the coefficients $A(\nu)$ of the eigenfunction continuum). The surprising result, which contradicts the conjecture, was that each of these components of the angular flux demonstrates a two-stage convergence: a relatively steep slope for the first few iterations, followed by a shallower slope for the remaining iterations. This is evidenced at both sides of the slab, as well as in the middle of the slab (see Figures 6–8, below).

When I showed these results to a colleague,[8] however, he recognized this pattern from his own realm of expertise. His specific remarks are in the Appendix.

Note that I have modified slightly the convergence metric. In Figures 6–8, I have plotted the magnitude of the *relative* residual scalar flux (*i.e.*, the residual scalar flux normalized to the final *total* computed scalar flux). This modification provides a direct visual indication for the precision of the calculation.

12. Analysis of the Flux Calculation Results

In the preceding sections, I have concentrated on the efficiency of the reduced-source calculations. In this section, I present evidence that the calculations are producing correct results.

Figure 9 presents the computed angular flux at three x-positions of the slab defined for the following calculations: $x = 0; \frac{1}{2}; 1$. The slab thickness is 1 mfp, and its 'material' specifications are as before: $c = 0.5$ (collision survival probability); analog capture; and isotropic scattering. The prescribed source at the plane of incidence ($x = 0$) is: $S(0, \mu) = \mu$ for $\mu > 0$.

Several prominent features of the plots in Figure 9 are reassuring. At $x = 0$, the flux for the inward-directed angles ($\mu > 0$) is virtually identically 1.0 (refer to Tables I and II, in Section 10 above), which is the analytic value. Similarly, at $x = 1$, the flux for the backward angles ($\mu < 0$) is virtually identically 0.0, which is also the analytic value. We note the expected discontinuities for $\mu = 0$ at both faces of the slab ($x = 0$ and $x = 1$), and the variation of the flux at the midpoint of the slab has characteristics one might expect from those that characterize the boundary values.

Aside from the agreement at the boundaries, as noted above, I was fortunate to get an independent benchmark for the slab interior, from Jerome Spanier [9] and Rong Kong.[10] Table III, below, compares our results. Mine were computed using the reduced-source method for 20 iterations. Spanier and Kong used an independent approach, being developed at the Claremont Graduate University.[9, 10]

TABLE III
Comparison of Reduced-Source Results and Results of Calculations at Claremont

Angular flux computed at $x = 0.5$ mfp		
Method	μ	$\Phi(0.5, \mu)$
Reduced Source	0.4995	0.4795008
"Claremont"	0.4995	0.4795063
Reduced Source	0.5005	0.4801207
"Claremont"	0.5005	0.4801261

The agreement, between the two independent calculational methods, to approximately 5 parts in $\frac{1}{2}$ million is very reassuring.

13. Summary

The reduced-source calculational approach for Monte Carlo transport in [x,μ] geometry has been demonstrated to produce highly efficient computational results. The efficiency in sample calculations was seen to exceed traditional figures-of-merit by multiple orders of magnitude.

The specific choice of sample problem was based on its two-dimensional character and the availability of a Monte Carlo technique for the kernel computation. These circumstances allowed the formulation of the present reduced-source investigation, whose aim was to demonstrate the far superior computing efficiency that obtains from such a formulation of an adaptive Monte Carlo technique.

The purpose here was not merely to offer an alternative solution to a problem that was solved a long time ago. Rather, this successful demonstration of exponential convergence for the sample problem comprises one milestone in an ongoing research effort at Los Alamos. The goal of the ongoing research is the advancement of adaptive Monte Carlo techniques so that applications, which are intractable with currently available methodology, will become practical. Convergence to the exact solution for the angular flux, concurrent with the convergence of the residual metric to zero, was seen to exhibit exponential rates, characterized by $(constant)^{-N}$ rather than the traditional Monte Carlo $\frac{constant}{\sqrt{N}}$, where N represents an appropriate unit of computing effort. An example of the latter is one iteration of a multi-iteration calculation (provided that each iteration takes approximately the same amount of computing time). The convergence metric used in these analyses was based on the magnitude of the residual scalar flux.

Acknowledgements

Consultations with Tom Booth and Jim Morel at Los Alamos National Laboratory were much appreciated, as were communications with Jerry Spanier and Rong Kong of Claremont Graduate University.

Appendix
Morel's Remarks Concerning Distinct Convergence Rates Evidenced by Boundary-Source Residual Calculations [8]

With reference to convergence characteristics observed and discussed in Section 11 of the main text, this type of convergence is reminiscent of the convergence of standard iterative schemes for solving linear systems of equations. Such iteration processes are analyzed in terms of the attenuation of errors by the iteration matrix. In particular, at the beginning of each iteration, the error is defined to be the current solution iterate subtracted from the exact solution. In many instances, this error can be expanded in terms of the eigenvectors of the iteration matrix. For instance, let the error after iteration m be given by:

$$\vec{E}^m = \sum_{n=1}^{N} a_n \vec{e}_n \qquad (A1)$$

where $\vec{e}_n$ is the n-th eigenvector of the iteration matrix, and a_n is its associated expansion coefficient. Each iteration is the functional equivalent of multiplying the error by the iteration matrix. Thus the error after iteration step $m+1$ is

$$\vec{E}^{m+1} = \sum_{n=1}^{N} a_n \lambda_n \vec{e}_n \qquad (A2)$$

where λ_n is the eigenvalue associated with $\vec{e}_n$. Every convergent iteration matrix has eigenvalues with magnitudes less than unity. Thus, as the iterations proceed, each eigenvector component of the error converges exponentially to zero at a rate proportional to the inverse of the magnitude of its associated eigenvalue. The actual convergence rate that one observes for the error as a whole can initially reflect the convergence rates of the most quickly converging eigenvectors, if these are dominant in the initial expansion of the error. But after a sufficient number of iterations, only the most slowly convergent eigenvector (or eigenvectors) will persist (all the others will have been attenuated to insignificance), and that slowest convergence rate will, thereafter, be observed for the error as a whole.

The use of residual equations plays a major role in the analysis of iteration methods for linear systems. This together with much empirical evidence suggests that there may be a deep connection between such iteration techniques and the boundary-source residual Monte Carlo technique.

FIGURE CAPTIONS

Figure 1. Decrease of (positive) reduced-source weight, as a function of iteration number, for a slab of thickness 2 [x-units]. The initial (incident) source density was of the form $S(\mu) = \mu e^{-\mu^2}$. Each iteration comprised 50 histories, and, for each of the runs (*i.e.*, using 1000 bins and 2000 bins) individually, the CPU-time per iteration was essentially constant. The plots compare how the refinement of the numerical computations affect the rate of decrease, and include a plot of $\frac{constant}{\sqrt{iteration}}$ for illustration.

Figure 2. Decrease of (negative) reduced-source weight, as a function of iteration number, for a slab of thickness 2 [x-units]. The initial (incident) source density was of the form $S(\mu) = \mu e^{-\mu^2}$. Each iteration comprised 50 histories, and, for each of the runs (*i.e.*, using 1000 bins and 2000 bins) individually, the CPU-time per iteration was essentially constant. The plots compare how the refinement of the numerical computations affect the rate of decrease, and include a plot of $\frac{constant}{\sqrt{iteration}}$ for illustration.

Figure 3. Convergence behavior of the magnitude of the residual scalar flux (integral of the residual angular flux) on the left boundary of a slab in $[x, \mu]$ geometry. The slab's 'material' specifications were: c=0.5 (collision survival probability); analog capture; and isotropic scattering. The x-dimension thickness was specified in units of mean-free-path [mfp]. The plot compares the results for 5 slab thicknesses: 0.02; 0.2; 2; 20; and 200. To improve the visual comparisons, the results were scaled by: 1 (unscaled), 0.1, 0.01, 0.001, and 0.0001, respectively.

Figure 4. Convergence behavior of the magnitude of the residual scalar flux (integral of the residual angular flux) on the right boundary of a slab in $[x, \mu]$ geometry. The slab's 'material' specifications were: c=0.5 (collision survival probability); analog capture; and isotropic scattering. The x-dimension thickness was specified in units of mean-free-path [mfp]. The plot compares the results for 5 slab thicknesses: 0.02; 0.2; 2; 20; and 200. To improve the visual comparisons, the results were scaled by: 1 (unscaled), 0.1, 0.01, 1, and 1, respectively.

Figure 5. Convergence behavior of the magnitude of the residual scalar flux (integral of the residual angular flux) at the left, middle, and right of a slab in $[x, \mu]$ geometry. The slab's 'material' specifications were: c=0.5 (collision survival probability); analog capture; and isotropic scattering. The x-dimension thickness was 40 mfp. The incident source (at the left plane of the slab) was $S(0, \mu) = \mu e^{-\mu^2}$.

Figure 6. Analysis of the convergence behavior, of the magnitude of the relative-residual scalar flux, for the 3 principal Case components, whose associated coefficients are denoted as $a_{\pm}$ and $A(\nu)$ (refer to Equation 1) in the main text). The $[x, \mu]$-geometry slab's 'material' specifications were: c=0.5 (collision survival probability); analog capture; and isotropic scattering. The x-dimension thickness was 20 mfp. The incident source (at the left plane of the slab) was $S(0, \mu) = \mu e^{-\mu^2}$. The behavior at the $x = 0$ plane of incidence is plotted.

Figure 7. Analysis of the convergence behavior, of the magnitude of the relative-residual scalar flux, for the 3 principal Case components, whose associated coefficients are denoted as $a_{\pm}$ and $A(\nu)$ (refer to Equation 1) in the main text). The $[x, \mu]$-geometry slab's 'material' specifications were: c=0.5 (collision survival probability); analog capture; and isotropic scattering. The x-dimension thickness was 20 mfp. The incident source (at the left plane of the slab) was $S(0, \mu) = \mu e^{-\mu^2}$. The behavior at the $x = 10$ slab middle is plotted.

Figure 8. Analysis of the convergence behavior, of the magnitude of the relative-residual scalar flux, for the 3 principal Case components, whose associated coefficients are denoted as $a_{\pm}$ and $A(\nu)$ (refer to Equation 1) in the main text). The $[x, \mu]$-geometry slab's 'material' specifications were: c=0.5 (collision survival probability); analog capture; and isotropic scattering. The x-dimension thickness was 20 mfp. The incident source (at the left plane of the slab) was $S(0, \mu) = \mu e^{-\mu^2}$. The behavior at the $x = 20$ exit plane is plotted.

Figure 9. Calculation of the angular flux for a slab in $[x, \mu]$ geometry. The slab's 'material' specifications were: c=0.5 (collision survival probability); analog capture; and isotropic scattering. The x-dimension thickness was 1 mfp. The incident source (at the $x = 0$ plane of the slab) was $S(0, \mu) = \mu$. The angular flux at the plane of incidence, $x = 0$, at the midplane, $x = 0.5$, and at the exit plane, $x = 1$, is compared.

Figure 1.

Figure 2.

Figure 3.

Figure 4.

Figure 5.

Figure 6.

Figure 7.

Figure 8.

Figure 9.

REFERENCES

[1] T. E. Booth, "Monte Carlo Estimates of Transport Solutions to the Isotropic Slab Problem," Los Alamos National Laboratory Report LA-UR-97-3871, Nuclear Science and Engineering: vol 130, Nov. 1998.

[2] Los Alamos National Laboratory Project, "Monte Carlo Methods for the 21^{st} Century." *http://www-xdiv.lanl.gov/XTM/projects/mc21/*

[3] Kenneth M. Case and Paul F. Zweifel, "Linear Transport Theory," Addison-Wesley (1967).

[4] George I. Bell and Samuel Glasstone, "Nuclear Reactor Theory," Van Nostrand Reinhold Company, Copyright 1970 by Litton Educational Publishing.

[5] H. Lichtenstein, "Convergence Rates for a Reduced-Source Monte Carlo Transport in $[x, \mu]$ Geometry," Los Alamos National Laboratory Report LA-UR-98-2172, X-Division Research Note XTM–RN (U) 98–014, to Distribution (May 13, 1998). *http://www.lanl.gov/home/hal/*

[6] Thomas E. Booth, Group X-CI, Project Principal Investigator, "Monte Carlo Methods for the 21^{st} Century," Los Alamos National Laboratory; private communications.

[7] J. H. Halton, J. Sci. Comp. **9**, 213 (1994).

[8] Jim Morel, Transport Methods Group Leader (X-TM), Los Alamos National Laboratory; private communications.

[9] Jerome Spanier, Joseph H. Pengilly Professor of Mathematics, Department of Mathematics, Claremont Graduate University, private communications.

[10] Rong Kong, Department of Mathematics, Claremont Graduate University, private communications.

HL:hal